The synergy of biology, intelligent systems, and space exploration

Eric Mjolsness and Ann Tavormina, Jet Propulsion Laboratory, California Institute of Technology

The farther we venture from home, the more self-reliant we need to be. This is as true in space exploration as it is in more commonplace aspects of the human experience.

In recent years, artificial intelligence and intelligent systems have enabled the initial demonstrations of autonomous capabilities on robotic exploration spacecraft. At the same time, intelligent systems have helped illuminate the means by which biological systems solve some of nature’s engineering challenges. It is often where disparate fields intersect that unexpected and useful insights into nature and engineering arise. And so it might be with intelligent systems, biology, and space exploration: our deepening understanding of biology is generating new ways to meet the challenges of space exploration and both fields are increasingly dependent on advances in computing. Directly and indirectly, intelligent systems can spur advances in space system engineering, letting us venture farther, perhaps sooner, into the solar system and nearby interstellar space.

The current decipherment of genetic information in biology is revolutionary. We stand at the edge of a deep understanding of biology at the system level, brought about by fundamental knowledge of mechanisms and strategies played out at the macromolecular level. Understanding in detail how biological systems store and retrieve information, control development, fabricate structural components, build molecular machines, sense the external environment, reproduce and disperse themselves throughout the environment, engage in error detection, and carry out self-repair can pay big dividends to space exploration. This type of knowledge will let us build more autonomous, more “self-reliant” space systems that can operate robustly in difficult environments farther and farther from Earth.

Here, we will discuss some of the ways that intelligent systems are advancing the understanding of human and other genomes as well as the molecular regulatory mechanisms for gene expression. We also extrapolate and speculate on space applications that could result from this new understanding of biology in such areas as human spaceflight (including space medicine and life support systems), NASA’s search for life elsewhere (astrobiology), and robotic spaceflight (biomanufacturing and in situ assembly).

Genomics

One of the most exciting things happening in biology is that a major community effort—the Human Genome Project—and similar projects for many other species are approaching their goals of acquiring nearly all the genetic information contained in the DNA of a living cell. How this 1D sequence of information affects the live functioning of a cell is also being revealed by a second wave of new instrumentation that quantifies the activity of thousands to tens of thousands of genes in a collection of cells, as well as other genome-wide clues to cellular function. The understanding of DNA sequence and cellular function information at the full genomic scale both depend on substantial computational effort. This effort includes special-purpose databases and pattern-recognition algorithms (central aspects of bioinformatics), as well as predictive dynamical simulations in computational biology. There is ample scope for intelligent systems to make further contributions to these emerging areas.

We already see major progress in sequence genomics and functional genomics revolutionizing the molecular study of cells in one species after another: numerous bacteria, and within the last two years baker’s yeast (Saccharomyces cerevisiae), a microscopic multicellular worm (C. elegans), a fruit fly (Drosophila melanogaster), and very soon human beings (Homo sapiens), to be followed by a more experimentally manipulable vertebrate, a mouse (Mus musculus) and at least one plant (Arabidopsis thaliana). In each case, we can expect major improvements in our understanding of how these organisms really work as molecular, cellular, and multicellular systems. We are just beginning to understand the technological implications of this cornucopia of new information.

Biology background: the central dogma and gene regulation

According to the central dogma of molecular biology, DNA is the heritable genetic material of the cell, which is transcribed into RNA, an informational intermediate, which is then translated into protein (see Figure 1). Usually, a specific gene (a heritable unit of information encoded in DNA) directs the production of a specific messenger RNA (mRNA) containing a subset of the information in the DNA. A special molecule called an RNA polymerase performs this transcription step. The mRNA in turn specifies a particular protein molecule that is created by the more elaborate process of translation, mainly performed by a large assembly or complex of molecules called the ribosome.
Transcription and translation together constitute gene expression: the action of a gene in producing protein molecules.

The proteins that result from gene expression are very diverse in their chemical and structural properties and can act as individual actuators, sensors, mechanical structure elements, motors, regulatory circuit elements, and catalysts in the cell. For example, a crucial protein in muscle is muscle myosin of the myosin-II family, which functions in a cyclic manner analogous to an automobile engine cylinder. We can model it as having several phases including a power stroke driven by energy from adenosine triphosphate (ATP) and a recovery stroke that readies it for the next power stroke. Each myosin molecule can “walk” along an appropriate substrate by itself if kept supplied with ATP for energy. In other words, myosin is a molecular motor. Very large numbers of these motor proteins are organized to act in concert for generating each and every one of our muscle movements—such as those required for breathing.

A very important and interesting class of proteins, the transcription factors, provides strong feedback from the realm of proteins to the genes by regulating the transcription, and thus the expression and action, of specific genes (see Figure 1). For example, MyoD is a transcription factor that regulates a number of other muscle-specific genes and proteins. It can trigger an entire program of cellular differentiation (irreversible specialization) to turn an immature cell in a developing mouse or human into part of a muscle cell. Transcription factors thus cast gene expression as a giant circuit of genes regulating other genes. With very few exceptions, mainly in the immune system, the circuit does not alter the gene itself, because it doesn’t change the genetic information encoded in the DNA molecule. The circuit simply affects what molecules are bound to the DNA and therefore how the gene is expressed in a particular cell. This gene-regulation circuit, composed of many interacting subcircuits, is a huge unclaimed prize for scientific understanding.

Thus each cell is an amazing, adaptive, and robust mechanical and informational system at the molecular level. The system acts under the coordination of genetic information that has probably evolved for robustness against many kinds of stochastic environmental challenges, among other selection criteria. Genes and their expression are the keys to understanding and controlling this system, for any terrestrial form of life.

An essential engineering aspect of molecular regulatory systems is their characteristic timescale. Transcriptional regulation of gene expression involves the slow (1,000 base pairs per minute) construction of long polymers (mRNA and proteins) and reflects relatively long-term decisions or other state information. Protein–protein interactions such as phosphorylation and noncovalent binding can also propagate regulatory information and are much faster (subseconds). Electrochemical neural information processing is even faster (milliseconds), yet still very slow compared to artificial electronics. The relatively slow timescales for molecular regulation are appropriate for regulating construction and fabrication operations.

Data from new kinds of instrumentation, including 2D microarrays (which measure mRNA gene expression in tens of thousands of genes simultaneously) and other new technologies (which work at the protein level) are now opening our eyes to detailed cellular control mechanisms governing the growth and development of biological organisms.

---

Richard Doyle, AI in Space Editor

---

**Figure 1.** Genetic information encoded in DNA is transcribed to messenger RNA (mRNA) by RNA polymerase II and related molecules. mRNA is translated to protein by a ribosome consisting of a complex or assemblage of multiple RNA and protein molecules (not shown). Some proteins (for example, MyoD in muscle) can then close the feedback loop by binding to DNA and regulating its transcription. Other proteins act as actuators, motors (such as muscle myosin), sensors, structural elements, and regulators of protein state. In this way, the genetic information in DNA acts as a “program” that controls the machinery and circuitry of the cell, including its own actions.
Intelligent systems

Intelligent systems are increasingly required for genomic data handling and analysis but also for hypothesis discovery and testing. Shotgun sequencing has accelerated the human and fruit fly genome projects and relies on large-scale computing to assemble many small DNA sequences into usable sequence data. This application has gone far to catapult bioinformatics into the critical path of mainstream biological science. Many Web-accessible sequence databases provide essential pattern recognition services for working biologists. Gene expression data is commonly analyzed using a variety of unsupervised learning techniques including k-means and other clustering algorithms, and self-organizing maps. The output of such analyses suggests useful clusters of related genes, introducing semi-automated hypothesis discovery as a major application of intelligent systems in biology.

In the short span of a few years, bioinformatics with intelligent systems capabilities has become critical to the development of biological science. Yet, the field is extremely dynamic and open to further computational contributions if well informed by biological collaborations.

A core technology for genomic biotechnology will be computational modeling and engineering at the cellular level. A challenging new analog of computer-aided circuit design will be required—the computer-aided design and optimization of life at the cellular level. In cell design, like circuit design, we must predictively model complex feedback networks, modelable to some degree as large sparse systems of ordinary differential equations with stochastic noise added, and optimize them to meet specifications. Unlike circuit design, however, we must begin by solving quite a difficult scientific problem: modeling aspects of an existing, complex biological regulatory network that has not been engineered for its understandability by human designers.

Fortunately, there are cell modeling and optimization methods that will work both for cell design and to fit the parameters in cell circuit models to biological data. We can use such cell optimization methods to fit genetic regulatory network models to gene expression data, or to design such models to meet engineering specifications. They must, however, be improved to work at the required genomic scale.

Because reproduction is one of the defining attributes of life, a natural place to start modeling and designing is with the cycle of cellular events involved in reproduction at the cellular level. Mathematical models are available for key aspects of cell-cycle regulation in a relatively simple eukaryotic species. Others have considered the more elaborate wiring of mammalian cell-cycle control as a circuit.

To meet the challenge of computational cell design, we will likely require many other aspects of intelligent systems technology. Intelligent text mining of scientific literature will also likely have a role, along with intelligent planning systems for laboratory experiments. Initially, we could use intelligent systems that model, design, and optimize cells on the ground to solve the space application problems we describe later. Eventually, however, the cell-biological analysis and design systems themselves might be placed onboard the spacecraft to meet the high demands for autonomous function in very ambitious deep space missions such as interstellar probes.

Space medicine

Let’s now look at possible applications in space of the confluence of computer science and biotechnology, as well as some of the future missions that might employ such capabilities.

Space medicine. Among space applications of genomics, it’s a relatively sure bet that space medicine will benefit greatly, enabling us to first understand and then treat some of the main health problems facing astronauts.

It is an inconvenient fact that astronauts’ health declines slowly but inexorably in space. Microgravity affects muscle and bone in ways that pose serious challenges for long-term expeditions to Mars and other off-planet destinations. For example, astronauts lose about 1% of bone mass per month spent in microgravity. Microgravity muscle atrophy occurs even faster. People are vulnerable to much higher space radiation hazards than on Earth, raising the specter of radiation sickness and cancer, especially in the pregnancies you would expect as part of long-term human exploration.

Happily, biotechnology is moving forward quickly to address many of these space health and medicine problems. The ability to measure and understand the action of expressed human genes in the body at the molecular level (what they do, not just what they say) through the proteins they make will revolutionize our ability to understand problems like microgravity-caused muscle atrophy and how to treat or prevent it.

Because genes are the key to all terrestrial life, we can expect enormous progress against those medical conditions that we choose to examine by the new gene expression technologies. Other breakthrough technologies are coming online as well; only within the past year or two, for example, has it become possible to systematically locate and measure gene expression in muscle satellite cells—the cells responsible for muscle regeneration following injury in mammals. Measuring gene expression in these cells opens up the possibility of understanding how they work in adult human beings. Microgravity muscle atrophy is one of a number of related muscle conditions and injuries that we should examine with these new technologies. Similar efforts are possible for microgravity bone loss and, with more difficulty, space radiation effects.

Clearly, a computational understanding of space-related muscle atrophy, bone loss, or cancers might have strong application to terrestrial medicine. Work in these areas (especially cancer) overlaps with mainstream biology and medicine and spin-offs are possible in either direction.

Astrobiology. The next application area is necessarily among the most speculative topics in science. The discoveries of planets in other solar systems, of a possible ocean under the ice on Europa, and of a history for water on Mars all bring the possibility of nonterrestrial life closer to scientific investigability. To pursue the possible discovery of such life, we’ll need to bring

Among space applications of genomics, it’s a relatively sure bet that space medicine will benefit greatly, enabling us to first understand and then treat some of the main health problems facing astronauts.
suitable instrumentation to bear on solar system sites and in telescopic spectral observations of extra-solar locations.

But what instrumentation or observation would be appropriate for detecting traces of unknown forms of life? Clues can be had from consideration of the great diversity of terrestrial life and habitats. Bacterial mats and biofilms are relatively macroscopic structures resulting from interacting behaviors of single-celled prokaryotic organisms.

On Earth, endolithic organisms leave observable chemical traces inside rocks, including some from deep underground. Visible spatio-chemical patterns in terrestrial rocks might arise from such biological origins but can also be produced by a surprising variety of abiotic pattern formation mechanisms. We will need detailed computational modeling of chemical and biological reaction networks to differentiate alternative biological and nonbiological hypotheses and subject them to discriminating observations. In this area, a computational, genome-wide understanding of terrestrial organisms would be quite valuable. It is believed that liquid water may have been sequestered under the surface of Mars. Understanding how endolithic organisms manifest in terrestrial samples will help us interpret drilled samples of rocks we hope to bring back someday from Mars for study on Earth.

In addition, a computational understanding of the evolution of genomic-scale regulatory networks will be essential to framing hypotheses about the origins of terrestrial RNA- and DNA-based life and of extraterrestrial life if it is presumed to possess some kind of molecular regulatory circuitry. In terrestrial life, entire regulatory networks are clearly related to one another evolutionarily; cross-species homologies are a major clue to understanding genomic data and the course of evolution.

Even if we find no nonterrestrial life in our solar system, preparing and executing the search will nevertheless advance the understanding of the origins of terrestrial life and will possibly enable us to reengineer terrestrial life for other locations as well as to certify that doing so won’t destroy indigenous life-forms elsewhere.

Biofabrication and morphogenesis. This application area might hold the broadest interest for previously nonbiological engineering disciplines. A computable understanding of cellular function is likely to lead to molecular engineering of particular biological mechanisms to solve engineering problems. Construction of wood, teeth, bones, hard shells, spider webs, and many other materials, minerals, and mechanical structures proceeds biologically by a process of evolutionarily optimized growth, starting with molecular building materials. The same is true, on a grander scale, for the construction of nervous systems and brains. Morphogenesis—the generation of multicellular form by signaling between cells that can grow, divide, and specialize—couples with molecular assembly between the cellular and subcellular scale to perform fabrication of organic, mineral, and computational structures in an efficient, adaptive and customizable way from elementary building units. Upon injury, aspects of the developmental self-fabrication process can be restarted to effect repair, resulting in a robust system. At least one bacterium, *Deinococcus radiodurans*, exhibits strong radiation resistance due perhaps to powerful genetic repair mechanisms that genomic studies are revealing now that it has been sequenced. These fabrication and repair capabilities are potentially valuable properties of a biofabrication technology governed by genetic and other regulatory networks.

For robotic space exploration missions, there are overwhelming reasons to learn from biofabrication and biological engineering. Perhaps chief among these reasons is the launch mass required to establish a self-sustaining industrial presence in space. Current space engineering projects are critically limited in scale and ambition by the requirement that their full mass be launched from Earth. A systematic alternative is to launch fabrication technology that can make use of mineral and material resources on low-gravity inner solar system bodies such as asteroids, moons, or even Mars. But the mass of manufacturing equipment is itself prohibitive when you consider all steps, including power generation, required to produce a useful item by conventional means. In principle biofabrication offers a very low-mass fabrication alternative, and so is the key to large-scale space engineering projects.

Because the nonterrestrial solar system is rich in mineral and energy resources, the ideal way to run a solar-system-wide economy might be to keep most of the people on Earth (where we enjoy the largest terrestrial environment that can be expected for centuries or millennia), beam up information such as industrial designs, plans, and encrypted command suites into a solar-system-wide industrial complex, and receive high-value finished goods through automated one-way descent transports manufactured in space. Earth would also receive massive return information feeds for science, engineering, entertainment, and so on. The solar system energy and materials mobilized by such infrastructure would also support big space projects such as creating a launcher for interstellar probes. This scenario minimizes the requirement for launch mass because few people need to be transported into space, but it would still require an enormous undertaking to lift whole industries (including construction equip-

Eric Mjolsness is a principal computer scientist at the Jet Propulsion Laboratory of the California Institute of Technology, and a Faculty Associate in Biology at the California Institute of Technology. His research interests include gene regulation networks, cell simulation, statistical pattern recognition, neural networks, computer vision, and large-scale optimization methods. He earned his AB in physics and mathematics from Washington University and his PhD in physics and computer science from the California Institute of Technology. He is a member of the IEEE and the ACM. Contact him at the Jet Propulsion Laboratory, MS 126-346, 4800 Oak Grove Drive, Pasadena CA 91109-8099; mjolsness@jpl.nasa.gov; www-aig.jpl.nasa.gov/mls.

Ann Tavormina is the manager of the Exploration Systems Autonomy section at JPL. Her research interests include developing technologies that can be incorporated in future spaceflight missions to advance the state of the art in space exploration. NASA’s growing programs in the search for life elsewhere and in astrobiology are of particularly keen interest. Formerly, she was the Mission Operations Manager of the Cassini Mission to Saturn and the Deputy Mission Director/Flight Engineering Office Manager of the Magellan Mission to Venus. She received a BS in biological sciences from Indiana University and a PhD in biophysics from the University of Pennsylvania. Contact her at the Jet Propulsion Laboratory, California Inst. of Technology, Mail Stop 126-347, 4800 Oak Grove Dr., Pasadena, CA 91109-8099; anna.m.tavormina@jpl.nasa.gov.
industry in space, at whatever level of human presence turns out to be most useful. And, of course, plants and oceanic algae produce oxygen from carbon dioxide, as is crucial to all animal life on Earth. If we can control plant development at the level of gene expression circuitry, as now seems very likely, we can modify plants to adapt to altered conditions in greenhouses throughout the solar system, and more importantly, to produce new chemical feedstocks required by local industry with a minimum of investment in launching massive capital equipment.

The biofabrication route to manufacturing has drawbacks as well. Living cells are not good metallurgists; they compute electrically with massive, slow ions rather than low-mass, fast electrons; and they have rotating motors at the molecular scale but not at larger scales. Blending biofabrication with current manufacturing technologies will present quite a challenge. One option is to imitate morphogenesis with larger electromechanical or robotic analogs of cells rather than reengineering actual living cells. This path essentially gives up on molecular-scale assembly and its adaptive and evolutionary optimization. Preferably, we can genetically engineer protein catalysts that perform unbiological solid-phase assembly tasks.

**Future space exploration mission concepts.** How we choose to capitalize on our growing understanding of biological mechanisms—directly, indirectly, or in a metaphorical sense—for the benefit of space exploration, depends on the particular problem being addressed.

Space missions thus enabled could include long-duration outposts in the solar system, with a high ratio of robotic to human occupants, as might be optimal for challenging space environments. Both AI and biotechnology could be brought to bear on the robotic as well as the human elements of the missions. Space medicine, astrobiology, biofabrication, and autonomy could have ready uses if the appropriate investments in technology development had been made and the results infused into mission plans early.

Exploration of the possible ocean under the ice of Jupiter’s moon Europa will require significant autonomy technology and can also benefit from advances in astrobiology. Exploration of Saturn’s moon Titan, which might have hydrocarbon oceans, could likewise benefit from these same technologies.

Demanding, long-term robotic missions to far outer solar system objects (such as in the Kuiper Belt or Oort Cloud) and interstellar missions would benefit greatly from the kind of radical self-repair and molecular-scale optimization technology that advances in biofabrication and autonomy are likely to provide.

As an extreme example of applying biofabrication to space exploration, consider the solar sail approach to interstellar spacecraft design (see Figure 2). It should be possible to get at least as far away from the Sun as the Kuiper Belt or Oort Cloud by using extremely lightweight and strong light-reflecting material surfaces actively controlled by a small low-power spacecraft over a long duration. Such a mission would require a high degree of autonomy and would be most effective if the sail and other components were self-repairing under the rain of cosmic radiation and molecularly optimized for their mechanical properties. Thus the solar sail could be “grown” using future biotechnology that incorporates the best properties of spider webs, flagella, and other biological structures.

One ambitious elaboration of the solar sail mission design would require high-power artificial directed radiation sources to accelerate the spacecraft. Such sources would be very substantial space-engineering projects, perhaps enabled by biofabri-
and intelligent systems is likely to redefine related mission concepts could lead to nearby stars, the synergy of biotechnology and potential payoffs of the synergy between engineering and evolved biological design, we can broaden and perhaps redirect our thinking about the essential technologies with which we will transform our world and explore the universe.

Long before we build probes that visit nearby stars, the synergy of biotechnology and intelligent systems is likely to redefine the technologies of space exploration. We have discussed potential applications in robotic and human spaceflight, astrobiology, and the development of the solar system. By considering the technical challenges and potential payoffs of the synergy between engineering and evolved biological design, we can broaden and perhaps redirect our thinking about the essential technologies with which we will transform our world and explore the universe.

References


